

Forwarding Deflection in Multi-Area OSPF

Stefano Iasi
Dipartimento di Automatica e
Informatica
Politecnico di Torino
Turin, Italy
stefano.iasi@studenti.polito.it

Pierre François
Computing Science and
Engineering Department
Université catholique de
Louvain
Louvain La Neuve, Belgium
pfr@info.ucl.ac.be

Steve Uhlig
Computing Science and
Engineering Department
Université catholique de
Louvain
Louvain La Neuve, Belgium
suh@info.ucl.ac.be

ABSTRACT

Multiple areas in OSPF have been introduced to improve the scalability of IGP routing in large networks. The principle is to partition the links of the network into *areas*, each link belonging to a single area. In this paper, we show that the way OSPF routers compute their routes makes it possible for *deflection* to arise, i.e. the forwarding path followed by IP packets differs from the routing path computed by the routing protocol. We give a necessary and sufficient condition for *deflection* to arise in a domain, and describe how to detect it. The goal is to provide a tool to decide where to establish OSPF virtual links to remove the deflection. The detection of deflection will be implemented in C-BGP [1], an open-source routing solver.

Categories and Subject Descriptors: [C.2.2]: Routing Protocols

General Terms: Algorithms, Design.

Keywords: OSPF, IGP routing, Forwarding.

1. INTRODUCTION

Hierarchical routing was introduced into OSPF to make it more scalable in large networks [2]. The principle is to divide the domain into several "areas". Each router has a detailed knowledge of the topology of the area(s) it belongs to. Flooding of Link State Advertisement (LSA) is used inside areas, while only summaries are exchanged across areas. A router may belong to several areas, in which case it is called a "border router" for these areas. A "border router" is a router such that at least two of its links belong to different areas. In (multi-area) OSPF, the network is hierarchically divided into a special area (called area 0) which makes the "backbone" of the OSPF network and all other areas. All border routers of the network must belong to this "backbone area", as well as to other areas.

A router reaches destinations in areas it does not belong to only through border routers of its own area. For that, each border router, let say belonging to area x , builds summary information about destinations it can reach in other areas. The IGP cost to reach a destination in another area is the cost towards the border router plus the cost that the border router advertises in its summary for this destination. When more than one border router can be used to reach a destination in another area, the route with the smallest total cost is chosen to reach the destination.

Here we shortly illustrate principles of best route computation in the context of multi-area OSPF. We consider only destinations

that belong to the OSPF domain, i.e. intra-area or inter-area destinations, not external destinations. An intra-area destination is a destination that lies in the same area as the considered router. Otherwise, a destination is an inter-area one.

Routes towards intra-area and inter-area destinations are computed differently. A router computes its intra-area routes by building a shortest-path tree based on the LSAs that it received from the routers of its area. Routes built in this way are tagged with an attribute whose value is set to "intra-area".

To build its routes towards inter-area destinations, a router examines the summaries injected by border routers inside its area. Only border routers are allowed to inject summaries inside their areas. As all border routers belong to area 0 (by definition of border routers in OSPF), they receive summaries from area 0 as well as other areas they belong to. In the computation of inter-area routes, border routers only consider summaries injected in area 0. For each inter-area destination announced, a router chooses among the possible routes the one having the smallest total IGP cost to reach the destination. The total cost of a route to reach a given destination is computed by adding to the IGP cost advertised by the border router in its summary, the IGP cost of the intra-area path to reach this border router. We call the border router used to exit the area the "egress router". An egress router must be reachable through an intra-area route. Routes built in this way are tagged with an attribute whose value is set to "inter-area".

2. ON FORWARDING DEFLECTION

In this section we explain how *deflection* can arise in multi-area OSPF. Two properties of OSPF inter-area routing (independently) lead to deflection.

Firstly, a border router BR_1 only considers summaries injected inside area 0. It will thus not be able to choose a path via an area $i \neq 0$ for a destination D if D is not in this area. However, BR_1 could be on the shortest path from a router X in area i to the border router BR_2 that X selected for destination D . When a packet with destination D , entering the network in X will reach BR_1 , it will be forwarded into the backbone, instead of being forwarded to BR_2 and only then be forwarded into the backbone. Packets can thus be forwarded out of a **non-backbone area** via a border router that is different from the border router that was selected by upstream routers.

Secondly, a router will always prefer an intra-area path over an inter-area path [2]. This means that if a Border Router BR_1 is connected to an area i , it will forward a packet towards a destination D in i along its shortest path to D within area i . However, BR_1 could be on the shortest path from a router X in the backbone to the border router BR_2 that X selected for destination D . When a packet

with destination D is in transit across the backbone and reaches X , it will be forwarded to BR_1 , and be deflected into area i , instead of being forwarded to BR_2 across the backbone and only then be forwarded into area i . The second property thus leads to packets being forwarded out of the **backbone** via a border router that is different from the border router that was selected by upstream routers.

3. EXAMPLE

Figure 1 illustrates an example of *deflection*. Figure 1 shows three different areas. The source of the traffic (S) is located in area 1, the destination (D) is located in area 2, and all border routers (denoted by BRx) belong to area 0 (backbone), as well as to other areas. BR_1 and BR_2 belong to area 1 and 0, and BR_3 belongs to area 2 and 0. Arrows on Figure 1 represent routing paths, not physical links.

We assume that all routers have computed their shortest paths towards all destinations, and that OSPF has converged inside the network. The routing table of each router contains all the routes for intra-area and inter-area destinations (see section 1).

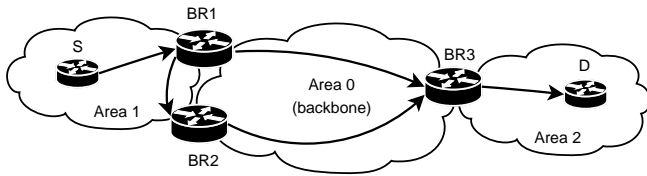


Figure 1: Example of deflection

Suppose that S wants to reach router D . The route that S uses towards D has BR_2 as its egress router. The intra-area path of S towards BR_2 crosses border router BR_1 . When a packet from S towards D arrives at BR_1 , the latter does not forward the packet to BR_2 on the path inside area 1 as thought by S . Rather, BR_1 sends the traffic towards D on one of its links belonging to area 0. Deflection arises in this example because S thinks that packets towards D will follow links within area 1 to reach BR_2 , while BR_1 will never use links belonging to area 1 to reach D . For BR_1 to reach D using links within area 1, D would have to lie inside area 1.

Even if the routing path chosen by S is optimal with respect to the routing information that it receives from the border routers of area 1, packets will be forwarded on a suboptimal path.

In this example, deflection takes place because BR_1 only considers summaries that are injected in the backbone, so that it will never forward packets towards D inside area 1.

4. DETECTING DEFLECTION

In this section, we explain how to detect deflection. Detecting a deflection should force an operator to establish an OSPF virtual link [2] to solve the problem¹.

We denote by $SP(X, Y)_i$ the shortest path from X to Y within area i . The cost of a path is simply the sum of the metrics of the links that it contains. $Adv(BR, Y)$ denotes the metric that is advertised for destination Y in the summary that is flooded by a border router BR .

Firstly, let us show how to detect deflection for packets forwarded from routers inside a non-backbone area i towards a destination D lying outside area i . If there is only one border router connected to

area i , no deflection can occur. Deflection will occur if there exists two border routers BR_1 and BR_2 such that

$$Cost(SP(BR_1, BR_2)_i) < Cost(SP(BR_1, BR_2)_0) \quad (1)$$

$$\exists D : Cost(SP(BR_1, BR_2)_i) + Adv(BR_2, D) < Adv(BR_1, D) \quad (2)$$

Note that if no router within area i uses BR_1 to reach BR_2 while the two conditions are met, routing will be sub-optimal but no deflection will occur. However, we did not make this condition mandatory as the usage of a virtual link is also recommended when sub-optimal routing is detected.

Secondly, let us show how to detect deflection for packets forwarded from routers inside the backbone towards a destination D lying in an area $i \neq 0$. Deflection will occur if there exists two border routers attached to area i , BR_1 and BR_2 , such that :

$$Cost(SP(BR_1, BR_2)_0) < Cost(SP(BR_1, BR_2)_i) \quad (1)$$

$$\exists D : Cost(SP(BR_1, BR_2)_0) + Adv(BR_2, D) < Adv(BR_1, D) \quad (2)$$

Note that, once again, if no router within that backbone uses BR_1 to reach BR_2 while the two conditions are met, routing will be sub-optimal but no deflection will occur. Indeed, in such a case, only BR_1 will forward packets directly into area i instead of choosing the optimal path via $SP(BR_1, BR_2)_0$.

To detect the presence of *deflection*, we rely on CBGP [1], an open-source routing solver designed to compute the state of the routing for large topologies containing many BGP routers. We are currently implementing inside CBGP the *deflection* check, which will also describe the characteristics of the virtual links that have to be established to solve the deflection.

5. CONCLUSIONS AND FURTHER WORK

In this paper we have studied *deflection* in OSPF, i.e. divergence between the routing path selected by a router towards a destination and the forwarding path followed by the traffic from this source to the destination. We have proposed necessary and sufficient conditions for *deflection* to occur inside multi-area OSPF.

Deflection as defined in this paper was limited to the case where OSPF has already converged. However, it is known that during the convergence of the IGP, forwarding loops may appear [3]. As further work, we will investigate how to detect deflection during the convergence of OSPF. We are currently implementing inside CBGP the *deflection* check, which will also describe the characteristics of the virtual links that have to be established to solve the deflection.

Acknowledgments

This work was partially supported by the E-NEXT NoE funded by the European Commission. Steve Uhlig is "chargé de recherches" FNRS (Fonds National de la Recherche Scientifique, Belgium).

6. REFERENCES

- [1] B. Quoitin, "C-BGP, an efficient BGP simulator," <http://cbgp.info.ucl.ac.be/>, September 2003.
- [2] J. Moy, "OSPF Version 2," Internet Engineering Task Force, RFC2328, April 1998.
- [3] P. Francois and O. Bonaventure, "Avoiding transient loops during IGP Convergence in IP Networks," in *Proc. of IEEE INFOCOM*, March 2005.

¹Due to space limitations, we do not show how virtual links are established.